BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 3: e37398
doi: 10.3897/biss.3.37398

OPEN ACCESS

Conference Abstract

PlutoF: Biodiversity data management platform for the complete data lifecycle

Urmas Kõljalg‡, Kessy Abarenkov§, Allan Zirk§, Veljo Runnel‡, Timo Piirmann‡, Raivo Pöhönen‡, Filipp Ivanov‡

‡ University of Tartu, Tartu, Estonia
§ University of Tartu, Natural History Museum, Tartu, Estonia

Corresponding author: Urmas Kõljalg (urmas.koljalg@ut.ee)

Received: 17 Jun 2019 | Published: 26 Jun 2019

Citation: Kõljalg U, Abarenkov K, Zirk A, Runnel V, Piirmann T, Pöhönen R, Ivanov F (2019) PlutoF: Biodiversity data management platform for the complete data lifecycle. Biodiversity Information Science and Standards 3: e37398. https://doi.org/10.3897/biss.3.37398

Abstract

PlutoF online platform (https://plutof.ut.ee) is built for the management of biodiversity data. The concept is to provide a common workbench where the full data lifecycle can be managed and support seamless data sharing between single users, workgroups and institutions. Today, large and sophisticated biodiversity datasets are increasingly developed and managed by international workgroups. PlutoF's ambition is to serve such collaborative projects as well as to provide data management services to single users, museum or private collections and research institutions.

Data management in PlutoF follows a logical order of the data lifecycle Fig. 1. At first, project metadata is uploaded including the project description, data management plan, participants, sampling areas, etc. Data upload and management activities then follow which is often linked to the internal data sharing. Some data analyses can be performed directly in the workbench or data can be exported in standard formats. PlutoF includes also data publishing module. Users can publish their data, generating a citable DOI without datasets leaving PlutoF workbench. PlutoF is part of the DataCite collaboration (https://datacite.org) and so far released more than 600 000 DOIs. Another option is to publish observation or collection datasets via the GBIF (Global Biodiversity Information Facility) portal. A. new feature implemented in 2019 allows users to publish High Throughput

© Kõljalg U et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Sequencing data as taxon occurrences in GBIF. There is an additional option to send specific datasets directly to the Pensoft online journals. Ultimately, PlutoF works as a data archive which completes the data life cycle.

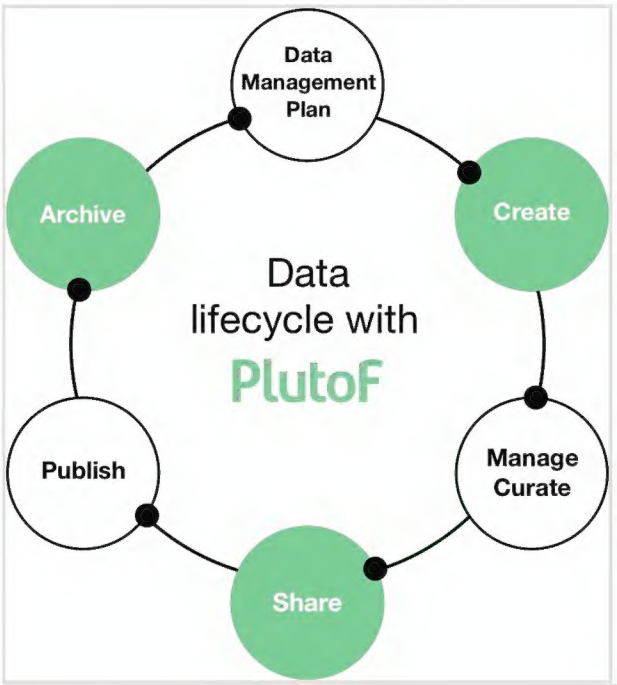

Figure 1.

Full data lifecycle with PlutoF platform. PlutoF follows standards, which ensure that users data are FAIR (Findable, Accessible, Interoperable, Reusable) throughout every phase of the data lifecycle.

In PlutoF users can manage different data types. Most common types include specimen and living specimen data, nucleotide sequences, human observations, material samples, taxonomic backbones and ecological data. Another important feature is that these data types can be managed as a single datasets or projects.

PlutoF follows several biodiversity standards. Examples include Darwin Core, GGBN (Global Genome Biodiversity Network), EML (Ecological Metadata Language), MCL (Microbiological Common Language), and MIxS (Minimum Information about any (x) Sequence).

Keywords

data management, biodiversity data, FAIR Data, data lifecycle

Presenting author

Allan Zirk

Presented at

Biodiversity_Next 2019

Funding program

PlutoF is funded by Estonian research infrastructure roadmap object NATARC (https://natarc.ut.ee/en/index.php). Future funding will be probably provided by the Est-DiSSCo which is listed in the new national roadmap. New business model will be created in 2020 which allows other institutions and countries to take part in the development of PlutoF.